

Claims

What is claimed is:

1. In a network, a method of processing received e-mail messages comprising:

- a) identifying information about an origin of a received e-mail message, the origin-identifying information based on data from at least one item of data in the message or at least one item of data used to send the message, the origin-identifying information including at least one of the following:
 - i) an actual sender of the message;
 - ii) a final IP address;
 - iii) a final domain name;
 - iv) a normalized reverse DNS lookup of the final IP address; or
 - v) an IP path used to send the message;
- b) categorizing whether the received message is solicited or unsolicited e-mail by using statistics based on information about the origin of the message; and
- c) processing the received message based on its categorization.

2. The method of claim 1 wherein the actual sender is identified by combining at least two of the following pieces of identifying information:

- a) an e-mail address used by a sender;
- b) a display name used by the sender;
- c) the final IP address used by the sender;
- d) the final domain name used by the sender;
- e) the normalized reverse DNS lookup of the final IP address;
- f) a mail from address;

- g) a domain name used by the sender;
- h) a name of client software used by the sender;
- i) user-agent;
- j) timezone;
- k) source IP address;
- l) sendmail version used by a first receiver;
- m) the IP path used to route the message;
- n) a range of IP addresses;
- o) owner data stored in databases; or
- p) a normalized mail from address.

3. The method of claim 1 further comprising using statistics compiled at at least one database to categorize whether the received message is solicited or unsolicited e-mail.

4. The method of claim 3 further comprising computing a score indicating a likelihood that the received message is unsolicited e-mail using statistics compiled at the at least one database.

5. The method of claim 3 further comprising sending information about received messages to the at least one database, the information including at least two of the following:

- a) information identifying the actual sender;
- b) whether the actual sender is included on a recipient's whitelist;
- c) whether the actual sender is included on a recipient's blacklist;

- d) information identifying the final IP address;
- e) whether the final IP address is included on the recipient's whitelist;
- f) whether the final IP address is included on the recipient's blacklist;
- g) information identifying the final domain name;
- h) whether the final domain name is included on the recipient's whitelist;
- i) whether the final domain name is included on the recipient's blacklist;
- j) information identifying the normalized reverse DNS lookup of the final IP address;
- k) whether the normalized reverse DNS lookup of the final IP address is included on the recipient's whitelist;
- l) whether the normalized reverse DNS lookup of the final IP address is included on the recipient's blacklist;
- m) information identifying the IP path;
- n) whether the IP path is included on the recipient's whitelist;
- o) whether the IP path is included on the recipient's blacklist;
- p) whether the message could be categorized locally; or
- q) whether a recipient changed a whitelist/blacklist status of the message.

6. The method of claim 3 further comprising requesting the at least one database to send a recipient statistics about origin-identifying information.

7. The method of claim 3 further comprising storing origin-identifying information from received messages including at least one of the following:

- a) a total number of messages sent;
- b) a number of messages sent over a first predetermined time period;
- c) a total number of messages sent to recipients in a network who have included the origin on a whitelist;
- d) a number of messages sent to recipient in the network who have included the origin on the whitelist over a second predetermined time period;
- e) a number of recipients who have included the origin on the whitelist;
- f) a total number of times a recipient changed the origin's whitelist/blacklist status;
- g) a number of times a recipient changed the origin's whitelist/blacklist status over a third predetermined time period;
- h) a total number of messages sent to recipients in the network who have not included the origin on the whitelist;
- i) a number of messages sent to recipients in the network who have not included the origin on the whitelist over a fourth predetermined time period;
- j) a total number of unique recipients in the network who have received at least one message from the origin;
- k) a total number of messages sent to unique recipients in a network who have included the origin on a whitelist; or
- l) a total number of messages sent to unique recipients in the network who have not included the origin on the whitelist.

8. The method of claim 3 wherein compiling statistics includes at least one of the following:

- a) determining a ratio of a first number of e-mail messages sent by the origin to recipients in the network who have included the origin on the whitelist in a predetermined time period divided by a second number of e-mail messages sent by the origin to recipients in the network in the predetermined time period;
- b) determining a ratio of a first number of recipients in the network who have included the origin on the whitelist divided by a second number of unique recipients in the network who received e-mails from the origin in the predetermined time period;
- c) determining a ratio of a first number of times in a predetermined time interval a message from the origin was moved from a whitelist to a blacklist divided by a second number of times a message from the origin was moved from a blacklist to a whitelist;
- d) determining a ratio of a first number of times in a predetermined time interval a message from the origin was moved from a blacklist to a whitelist divided by a second number of times a message from the origin was moved from a whitelist to a blacklist;
- e) determining a ratio of a first number of unique users within the network who whitelisted the origin within a predetermined time period compared to a second number of unique users within the network who blacklisted the origin within the predetermined time period;
- f) determining a ratio reflecting whether the origin sends a majority of messages to recipients who have included the origin on the whitelist;
- g) determining a ratio reflecting a first number of wanted messages sent by the origin compared to a second number of unwanted or total messages sent by the origin;

- h) determining a difference between a first number of expected messages sent by the origin and a second number of unexpected messages sent by the origin;
- i) determining a difference between a first number of times a user whitelisted a message from the origin and a second number of times a user blacklisted a message from the origin;
- j) determining a difference reflecting whether the origin sends a majority of messages to known recipients;
- k) converting any of the above ratios or differences to a score indicating the likelihood the message is unsolicited e-mail; or
- l) applying the score to the appropriate messages in the spam folder.

9. The method of claim 3 further comprising setting a predetermined threshold for accepting messages based on statistics associated with the origin-identifying information.

10. A computer-readable storage medium having instructions which, when read, cause a computer to perform a method of processing received e-mail messages comprising:

- a) identifying information about an origin of a received e-mail message, the origin-identifying information based on data from at least one item of data in the message or at least one item of data used to send the message, the origin-identifying information including at least one of the following:
 - i) an actual sender of the message;
 - ii) a final IP address;
 - iii) a final domain name;

- iv) a normalized reverse DNS lookup of the final IP address; or
- v) an IP path used to send the message;
- b) categorizing whether the received message is solicited or unsolicited e-mail by using statistics based on information about the origin of the message; and
- c) processing the received message based on its categorization.

11. The computer-readable storage medium of claim 10 wherein the actual sender is identified by combining at least two of the following pieces of identifying information:

- a) an e-mail address used by a sender;
- b) a display name used by the sender;
- c) the final IP address used by the sender;
- d) the final domain name used by the sender;
- e) the normalized reverse DNS lookup of the final IP address;
- f) a mail from address;
- g) a domain name used by the sender;
- h) a name of client software used by the sender;
- i) user-agent;
- j) timezone;
- k) source IP address;
- l) sendmail version used by a first receiver;
- m) the IP path used to route the message;
- n) a range of IP addresses;
- o) owner data stored in databases; or
- p) a normalized mail from address.

12. The computer-readable storage medium of claim 10 the method further comprising using statistics compiled at at least one database to categorize whether the received message is unsolicited e-mail.

13. The computer-readable storage medium of claim 12, the method further comprising using statistics compiled at the at least one database to compute a score indicating a likelihood that the received message is solicited or unsolicited e-mail.

14. The computer-readable storage medium of claim 12, the method further comprising sending information about received messages to the at least one database, the information including at least two of the following:

- a) information identifying the actual sender;
- b) whether the actual sender is included on a recipient's whitelist;
- c) whether the actual sender is included on a recipient's blacklist;
- d) information identifying the final IP address;
- e) whether the final IP address is included on the recipient's whitelist;
- f) whether the final IP address is included on the recipient's blacklist;
- g) information identifying the final domain name;
- h) whether the final domain name is included on the recipient's whitelist;
- i) whether the final domain name is included on the recipient's blacklist;
- j) information identifying the normalized reverse DNS lookup of the final IP address;

- k) whether the normalized reverse DNS lookup of the final IP address is included on the recipient's whitelist;
- l) whether the normalized reverse DNS lookup of the final IP address is included on the recipient's blacklist;
- m) information identifying the IP path;
- n) whether the IP path is included on the recipient's whitelist;
- o) whether the IP path is included on the recipient's blacklist;
- p) whether the message could be categorized locally; or
- q) whether a recipient changed a whitelist/blacklist status of the message.

15. The computer-readable storage medium of claim 12, the method further comprising requesting the at least one database to send a recipient statistics about origin-identifying information.

16. The computer-readable storage medium of claim 12, the method further comprising storing origin-identifying information from received messages including at least one of the following:

- a) a total number of messages sent;
- b) a number of messages sent over a first predetermined time period;
- c) a total number of messages sent to recipients in a network who have included the origin on a whitelist;
- d) a number of messages sent to recipient in the network who have included the origin on the whitelist over a second predetermined time period;
- e) a number of recipients who have included the origin on the whitelist;
- f) a total number of times a recipient changed the origin's whitelist/blacklist status;
- g) a number of times a recipient changed the origin's whitelist/blacklist status over a third predetermined time period;
- h) a total number of messages sent to recipients in the network who have not included the origin on the whitelist;
- i) a number of messages sent to recipients in the network who have not included the origin on the whitelist over a fourth predetermined time period;
- j) a total number of unique recipients in the network who have received at least one message from the origin;
- k) a total number of messages sent to unique recipients in a network who have included the origin on a whitelist; or
- l) a total number of messages sent to unique recipients in the network who have not included the origin on the whitelist.

17. The computer-readable storage medium of claim 12 wherein compiling statistics includes at least one of the following:

- a) determining a ratio of a first number of e-mail messages sent by the origin to recipients in the network who have included the origin on the whitelist in a predetermined time period divided by a second number of e-mail messages sent by the origin to recipients in the network in the predetermined time period;
- b) determining a ratio of a first number of recipients in the network who have included the origin on the whitelist divided by a second number of unique recipients in the network who received e-mails from the origin in the predetermined time period;
- c) determining a ratio of a first number of times in a predetermined time interval a message from the origin was moved from a whitelist to a blacklist divided by a second number of times a message from the origin was moved from a blacklist to a whitelist;
- d) determining a ratio of a first number of times in a predetermined time interval a message from the origin was moved from a blacklist to a whitelist divided by a second number of times a message from the origin was moved from a blacklist to a whitelist;
- e) determining a ratio of a first number of unique users within the network who whitelisted the origin within a predetermined time period compared to a second number of unique users within the network who blacklisted the origin within the predetermined time period;
- f) determining a ratio reflecting whether the origin sends a majority of messages to recipients who have included the origin on the whitelist;

- g) determining a ratio reflecting a first number of wanted messages sent by the origin compared to a second number of unwanted or total messages sent by the origin;
- h) determining a difference between a first number of expected messages sent by the origin and a second number of unexpected messages sent by the origin;
- i) determining a difference between a first number of times a user whitelisted a message from the origin and a second number of times a user blacklisted a message from the origin;
- j) determining a difference reflecting whether the origin sends a majority of messages to known recipients;
- k) converting any of the above ratios or differences to a score indicating the likelihood the message is unsolicited e-mail; or
- l) applying the score to the appropriate messages in the spam folder.

18. The computer-readable storage medium of claim 12, the method further comprising setting a predetermined threshold for accepting messages based on statistics associated with the origin-identifying information.

19. In a network, a method for categorizing received e-mail messages comprising:

- a) collecting information about an origin of an e-mail message, the origin-identifying information including at least one of the following:
 - i) an actual sender;
 - ii) a final IP address;
 - iii) a final domain name;

- iv) a normalized reverse DNS lookup of the final IP address; or
- v) an IP path used to send the message; and
- b) compiling statistics at at least one database about the origin based on the origin-identifying information, wherein the statistics are used to categorize whether the received message is solicited or unsolicited.

20. The method of claim 19 wherein the actual sender is identified by combining at least two of the following pieces of identifying information:

- a) an e-mail address used by a sender;
- b) a display name used by the sender;
- c) the final IP address used by the sender;
- d) the final domain name used by the sender;
- e) the normalized reverse DNS lookup of the final IP address;
- f) a mail from address;
- g) a domain name used by the sender;
- h) a name of client software used by the sender;
- i) user-agent;
- j) timezone;
- k) source IP address;
- l) sendmail version used by a first receiver;
- m) the IP path used to route the message;
- n) a range of IP addresses;
- o) owner data stored in databases; or
- p) a normalized mail from address.

21. The method of claim 19 further comprising using statistics compiled at the at least one database to compute a score indicating a likelihood that the received message is solicited or unsolicited e-mail.

22. The method of claim 19 further comprising sending information about received messages to the at least one database, the information including at least two of the following:

- a) information identifying the actual sender;
- b) whether the actual sender is included on a recipient's whitelist;
- c) whether the actual sender is included on a recipient's blacklist;
- d) information identifying the final IP address;
- e) whether the final IP address is included on the recipient's whitelist;
- f) whether the final IP address is included on the recipient's blacklist;
- g) information identifying the final domain name;
- h) whether the final domain name is included on the recipient's whitelist;
- i) whether the final domain name is included on the recipient's blacklist;
- j) information identifying the normalized reverse DNS lookup of the final IP address;
- k) whether the normalized reverse DNS lookup of the final IP address is included on the recipient's whitelist;
- l) whether the normalized reverse DNS lookup of the final IP address is included on the recipient's blacklist;
- m) information identifying the IP path;

- n) whether the IP path is included on the recipient's whitelist;
- o) whether the IP path is included on the recipient's blacklist;
- p) whether the message could be categorized locally; or
- q) whether a recipient changed a whitelist/blacklist status of the message.

23. The method of claim 19 further comprising requesting the at least one database to send a recipient statistics about origin-identifying information.

24. The method of claim 19 further comprising storing origin-identifying information about messages including at least one of the following:

- a) a total number of messages sent;
- b) a number of messages sent over a first predetermined time period;
- c) a total number of messages sent to recipients in a network who have included the origin on a whitelist;
- d) a number of messages sent to recipient in the network who have included the origin on the whitelist over a second predetermined time period;
- e) a number of recipients who have included the origin on the whitelist;
- f) a total number of times a recipient changed the origin's whitelist/blacklist status;
- g) a number of times a recipient changed the origin's whitelist/blacklist status over a third predetermined time period;

- h) a total number of messages sent to recipients in the network who have not included the origin on the whitelist;
- i) a number of messages sent to recipients in the network who have not included the origin on the whitelist over a fourth predetermined time period;
- j) a total number of unique recipients in the network who have received at least one message from the origin;
- k) a total number of messages sent to unique recipients in a network who have included the origin on a whitelist; or
- l) a total number of messages sent to unique recipients in the network who have not included the origin on the whitelist.

25. The method of claim 19 wherein compiling statistics includes at least one of the following:

- a) determining a ratio of a first number of e-mail messages sent by the origin to recipients in the network who have included the origin on the whitelist in a predetermined time period divided by a second number of e-mail messages sent by the origin to recipients in the network in the predetermined time period;
- b) determining a ratio of a first number of recipients in the network who have included the origin on the whitelist divided by a second number of unique recipients in the network who received e-mails from the origin in the predetermined time period;
- c) determining a ratio of a first number of times in a predetermined time interval a message from the origin was moved from a whitelist to a blacklist divided by a second number of times a message from the origin was moved from a blacklist to a whitelist;

- d) determining a ratio of a first number of times in a predetermined time interval a message from the origin was moved from a blacklist to a whitelist divided by a second number of times a message from the origin was moved from a blacklist to a whitelist;
- e) determining a ratio of a first number of unique users within the network who whitelisted the origin within a predetermined time period compared to a second number of unique users within the network who blacklisted the origin within the predetermined time period;
- f) determining a ratio reflecting whether the origin sends a majority of messages to recipients who have included the origin on the whitelist;
- g) determining a ratio reflecting a first number of wanted messages sent by the origin compared to a second number of unwanted or total messages sent by the origin;
- h) determining a difference between a first number of expected messages sent by the origin and a second number of unexpected messages sent by the origin;
- i) determining a difference between a first number of times a user whitelisted a message from the origin and a second number of times a user blacklisted a message from the origin;
- j) determining a difference reflecting whether the origin sends a majority of messages to known recipients;
- k) converting any of the above ratios or differences to a score indicating the likelihood the message is unsolicited e-mail; or
- l) applying the score to the appropriate messages in the spam folder.

26. The method of claim 19 further comprising setting a predetermined threshold for accepting messages based on statistics associated with origin-identifying information.